

**PCT**

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
 International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>C12N 15/00, 15/64, 15/66, C12P 19/34, 21/06, C12Q 1/68</b>		<b>A1</b>	(11) International Publication Number: <b>WO 00/12687</b>
			(43) International Publication Date: 9 March 2000 (09.03.00)
(21) International Application Number: <b>PCT/US99/19413</b>		(74) Agent: REITER, Stephen, E.; Gray Cary Ware & Freidenrich LLP, Suite 1600, 4365 Executive Drive, San Diego, CA 92121 (US).	
(22) International Filing Date: 25 August 1999 (25.08.99)			
(30) Priority Data: 09/141,935 28 August 1998 (28.08.98) US		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(63) Related by Continuation (CON) or Continuation-In-Part (CIP) to Earlier Application US 09/141,935 (CON) Filed on 28 August 1998 (28.08.98)			
(71) Applicant (for all designated States except US): INVITROGEN CORPORATION [US/US]; 1600 Faraday Avenue, Carlsbad, CA 92008 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): MILES, David, J. [US/US]; 2027 Haller Street, San Diego, CA 92104 (US). TURNER, Lyle, C. [US/US]; 880 Munevar Road, Cardiff, CA 92007 (US). MARCIL, Robert [US/US]; 1600 Faraday Avenue, Carlsbad, CA 92008 (US). MC CONNELL, Gina, C. [US/US]; 13952 Stoney Gate Place, San Diego, CA 92128 (US).		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: SYSTEM FOR THE RAPID MANIPULATION OF NUCLEIC ACID SEQUENCES			
(57) Abstract  The present invention is a cell-free subcloning system utilizing three elements: (1) a donor vector that contains a nucleic acid sequence to be transferred to another vector flanked by a site-specific recombination sequence and one or more optional additional nucleic acid sequences, (2) an acceptor vector that contains a site-specific recombination sequence and one or more optional additional nucleic acid sequences, and (3) a site-specific recombinase that recognizes the site-specific recombination sequences in the donor and acceptor vectors so as to transfer the transfer sequence from the donor to the acceptor vector upon contact of the three elements of the system. Also disclosed are rapid subcloning methods employing the vectors and enzymes disclosed herein and kits for use in such methods.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## Description

5

10

15

20

25

30

35

40

45

50

55

**SYSTEM FOR THE RAPID MANIPULATION  
OF NUCLEIC ACID SEQUENCES**

**Field of the Invention**

5           The invention disclosed herein relates to the field of molecular biology and methods useful therefor. More particularly the invention relates to methods for subcloning of nucleic acid sequences.

**Background of the Invention**

10           The discovery and isolation of restriction endonucleases, specific enzymes  
15           capable of manipulating nucleic acid sequences, precipitated a revolution in molecular biological techniques. Restriction endonucleases were used to cut large DNAs into smaller fragments that could be re-attached to heterologous pieces of DNA by ligases. These techniques allowed scientists to transfer a gene encoding a particular protein into a relatively small plasmid vector that could be transfected into a cell for  
20           production of the encoded protein.

25           Over the years, a large number of vectors have been developed for a wide variety of specialized research, manufacturing, and production uses. For example, many types of expression vectors have been developed that allow heterologous proteins to be expressed in an increasingly larger number of cell types, including  
30           insect, plant, mammalian, and bacterial cells. Among expression vectors, specialized vectors have been developed that facilitate large scale production of proteins, for instance, by increasing levels of the protein produced or by introducing elements into the protein that aid in purification. Other vectors have been designed for use in specific research protocols, such as conducting one-hybrid or two-hybrid screens.  
35           Each specialized vector contains a specific set of nucleic acid sequences that give it its particular features. No one vector can contain all of these features, however, as the vector would eventually become too large to be easily manipulated. Thus, a nucleic acid sequence of interest must be moved from one vector to another as different specialized needs arise, a process known as subcloning.

Conventional subcloning methods require that each vector into which a nucleic acid sequence is to be subcloned contain restriction endonuclease recognition/digestion sites that are absent in the nucleic acid sequence in order to prevent the nucleic acid sequence from being cut into one or more pieces when subjected to the restriction endonuclease for removal from the vector and passage to the next vector. One must, therefore, either know the entire sequence of the nucleic acid being subcloned or test it with each restriction endonuclease proposed for use to see if it contains a matching recognition site. Either process requires time and resources to perform.

In addition, conventional subcloning methods require that the nucleic acid sequence being subcloned have sequences at its 5' and 3' ends that match the restriction endonuclease site into which it is being inserted. As not all available vectors have the same restriction endonuclease sites, the nucleic acid sequence to be transferred must usually be modified at its ends to make it compatible with each vector to be used in subcloning techniques.

Another drawback to conventional subcloning techniques is the use of ligases. These enzymes are relatively slow acting, require ATP, and generally are highly temperature sensitive

A need still exists in the art, therefore, for a simple, rapid system for the manipulation of nucleic acid sequences between vectors. The present invention addresses that need.

#### **Brief Description of the Invention**

The present invention comprises a cell-free subcloning system, methods for the rapid manipulation and subcloning of nucleic acid sequences using the system, and kits suitable for use in conducting such methods. In general, the system, methods, and kits of the invention utilize three elements. The first element is a donor vector comprising (1) a transfer sequence of nucleic acid to be transferred to an acceptor

5 vector, (2) a site specific recombination nucleic acid sequence flanking the transfer  
sequence as shown in Figure 1A; and, (3) optionally, one or more additional nucleic  
10 acid sequences. The second element is an acceptor vector comprising (1) a site-  
specific recombination sequence that matches the site-specific recombination  
5 sequence of the donor vector as shown in Figure 1B, and (2) one or more additional  
nucleic acid sequences. The third element is a site-specific, ATP independent  
15 recombinase, that recognizes the site specific recombination sequences in both the  
donor and acceptor vectors.

20 The site-specific recombinases employed in the practice of the present  
invention are enzymes that spontaneously recognize and cleave at least one strand of a  
double strand of nucleic acids within a sequence segment known as the site-specific  
25 recombination sequence. In the donor vector, the site specific recombination  
sequences are placed contiguously on either side of (i.e., "flank") a transfer sequence  
of nucleic acid whose excision from the donor vector and transfer to the acceptor  
15 vector is desired. In use, the donor vector containing the transfer sequence and the  
acceptor vector are placed within a single cell-free solution. Upon addition of the site-  
specific recombinase to the cell-free solution, the transfer sequence is excised from  
30 the donor vector. In some portion of the acceptor vectors in the cell-free solution (i.e.,  
"occasionally") the excised transfer sequence is ligated into the acceptor vector by  
35 operation of the recombinase upon the site-specific recombination sequence, without  
the use of a separate ligase to accomplish the ligation. The acceptor vectors generally  
further comprise a selectable marker gene to aid in identifying and isolating from the  
40 cell-free solution using known methods those acceptor vectors into which the transfer  
sequence has been successfully inserted. The site-specific recombination sequences  
25 of the donor and acceptor vehicles are preferably identical, but can vary in nucleic  
acid sequence so long as recognition of the site-specific recombination sequence by  
45 the recombinase is preserved despite the variance.

50 The present invention thus affords a novel single-step method and associated  
vectors and kits for moving nucleic acid sequences, such as recombinant DNA  
30 molecules, from one type of subcloning vector to another that overcomes the above-

described problems in the art. For example, the invention eliminates the need for incorporation of "add on" base sequences to transfer sequence to provide unique restriction sites.

In particular, topoisomerase-based cloning circumvents any problems associated with addition of nontemplated nucleotides by DNA polymerase at the 3' end of the amplified DNA. Any nontemplated base (N) at the 3' end of a PCR product destined for topoisomerase-based transfer (GCCCTTxxxxN-3') will dissociate spontaneously upon covalent adduct formation, and will therefore have no impact on the ligation to vector. Second, the only molecule that can possibly be ligated into the acceptor vector is the covalently activated transfer sequence and the transfer sequence can only be transferred to the acceptor vector. There is no potential for in vitro covalent closure of the acceptor vector itself, which ensures low background. There is also no opportunity for the transfer sequences to ligate to one another, which precludes cloning of concatameric repeats. In addition, unintended internal restriction of an uncharacterized sequence is avoided because the use of common restriction enzymes is avoided.

#### Description of the Figures

FIGURE 1A is a double stranded nucleic acid sequence (SEQ ID NO:17 and complementary strand thereto) representing a donor vector with a double stranded nucleic acid transfer sequence flanked by topoisomerase I recombinase recognition sites (single underlined) with a 4 base core sequence (within brackets).

FIGURE 1B is a double stranded nucleic acid (SEQ ID NO:18 and complementary strand thereto) representing an invention acceptor vector containing two recombinase recognition sites that match those in the donor vector and 10 base pair spacer sequences (double underlined) ready to receive a transfer sequence.

FIGURE 1C is a double stranded nucleic acid (SEQ ID NO:19 and complementary strand thereto) representing a new recombinant vector created by the operation of topoisomerase I upon the donor and acceptor vectors of Figure 1A and 1B, respectively. The transfer sequence is now inserted into the acceptor vector.

FIGURE 2 is schematic representation of the method of the invention utilizing a donor vector ("pDonor") containing a selectable marker gene other than Zeocin, an origin of replication sequence ("ori"), and a transfer sequence ("gene of interest") flanked by lox P recognition sites. The acceptor vector ("pAcceptor") contains a gene encoding resistance to the antibiotic Zeocin™ ("Zeo"), an origin of replication sequence ("ori"), and a gene encoding ccdB, a lethal compound, flanked by loxP sites. The arrow indicates that when the donor and acceptor vectors are combined in a reaction mixture in the presence of the recombinase Cre, a new recombinant vector ("pRecombinant") is created, which recombinant vector contains the transfer sequence and a gene encoding Zeo. Cells transformed with the reaction mixture will grow in the presence of the antibiotic Zeocin™ only if the recombination event has successfully occurred.

#### Detailed Description of the Invention

In one embodiment of the invention, there is provided a cell-free subcloning system comprising (1) a donor vector comprising a transfer sequence flanked by site-specific recombination sequences, (2) an acceptor vector comprising a site-specific recombination sequence that matches the site-specific recombination sequences of the donor vector, and (3) a site-specific recombinase capable of recognizing the site-specific recombination sequence. Each vector is of duplex nucleic acid sequence, and the transfer is a bivalent strand transfer. In the acceptor vector, the transfer sequence will be inserted in the immediate vicinity of and downstream of, or can be adjacent to, the site-specific recombination sequence. Optionally, one or more additional nucleic acid sequences, such as a selection marker gene, an origin of replication, a promoter-enhancer sequence, and the like can be included in the donor and acceptor vectors.



5 The subcloning event occurs in a cell-free environment without the need to use  
restriction enzyme(s), and the transfer of the transfer sequence to the acceptor vector  
occurs without the expense of ATP.

10 In a presently preferred embodiment of the invention, following the site-  
5 specific recombination event that occurs between the site-specific recombination  
sequences located on each vector (i.e., the donor and acceptor vectors), the transfer  
15 sequence is inserted into the acceptor vector in a manner that retains the proper  
translational reading frame of the transfer sequence.

20 As used herein "vector" means a recombinant nucleic acid sequence of duplex  
10 DNA that has been constructed to comprise one or more functional units not found  
together in nature. Examples include circular, double-stranded, extrachromosomal  
DNA molecules (plasmids), cosmids (plasmids containing COS sequences from  
25 lambda phage), viral genomes comprising non-native nucleic acid sequences, and the  
like. When used in the context of describing a vector, the terms "donor" and  
15 "acceptor" refer to the fact that one vector (the donor) will contain a nucleic acid  
sequence, referred to herein as the "transfer sequence," that is to be excised and  
30 transferred to another (the acceptor) vector. Any given vector can be a donor or an  
acceptor, depending on whether it is the vector from which a nucleic acid sequence is  
being transferred, or the vector into which a nucleic acid sequence is introduced.

35 Both donor and acceptor vectors contain site-specific recombination  
20 sequences, which are sequences of nucleic acids that are specifically recognized by a  
particular site-specific recombinase. Site specific recombinases, as the term is used  
40 herein, are enzymes that catalyze the excision and/or recombination of nucleic acid  
sequences, and may form intermediate complexes with the transfer sequence DNA  
25 during the recombination event. These enzymes recognize a relatively short, unique  
nucleic acid sequence in the donor and acceptor vectors that serves as a site for both  
45 recognition and recombination. Recombinases particularly useful in the practice of  
the invention are those that function in a wide variety of cell types because such  
enzymes do not require any host specific factors and do not require ATP to function.

5 Examples of site-specific recombinases of this type include type I topoisomerases (S.  
Shuman, *J. Biological Chemistry* 266:11372-79, 1991), integrases (Argos, *et al.*,  
10 *EMBO J* 5:433-440, 1986), resolvases (Hallet and Sherratt, *FEMS Microbiol. Rev.*  
21:157-178, 1997), and the like.

5 A particularly suitable enzyme for use in the practice of the invention is a type  
I topoisomerase, particularly vaccinia DNA topoisomerase. Vaccinia DNA  
15 topoisomerase binds to duplex DNA and cleaves the phosphodiester backbone of one  
strand. The enzyme exhibits a high level of sequence specificity, akin to that of a  
restriction endonuclease. Cleavage preferentially occurs at a consensus  
20 10 pentapyrimidine element 5'-(C/T)CCTT↓ (SEQ ID NO: 1) in the scissile strand. In  
the cleavage reaction, bond energy is conserved via the formation of a covalent adduct  
between the 3' phosphate of the incised strand and a tyrosyl residue of the  
25 topoisomerase I protein. Vaccinia topoisomerase can religate the covalently held  
strand across the same bond originally cleaved (as occurs during DNA relaxation) or it  
15 can ligate the strand to a heterologous acceptor DNA 5' end containing a site specific  
recombination site, such as the DNA in the invention acceptor vector, and thereby  
30 create a new recombinant molecule, as shown in Figure 1C.

When the substrate is configured such that the scissile bond with the  
topoisomerase is situated near (within about 10 to about 12 base pairs of) the 3' end of  
35 20 a DNA duplex, cleavage is accompanied by the spontaneous dissociation of the  
downstream portion of the cleaved strand in the donor vector. The resulting  
topoisomerase-DNA complex, containing a 5' single-stranded tail, can religate to an  
40 acceptor DNA if the acceptor molecule has a 5' OH terminated acceptor strand with  
sequence (e.g. of at least a four base overhang) complementary to that of the activated  
25 donor complex (i.e., the single strand tail of the noncleaved donor strand in the  
immediate vicinity of the scissile phosphate). In the absence of an acceptor strand, the  
45 topoisomerase can transfer the CCCTT strand to water, releasing a 3'-phosphate-  
terminated hydrolysis product, or to glycerol. However, the hydrolysis reaction is  
much slower than religation to an acceptor DNA strand of the acceptor vector, the  
50 30 extent of strand transfer to non-DNA nucleophiles being generally about 14-40%.

5 The specificity of vaccinia topoisomerase in DNA cleavage and its versatility  
in strand transfer have inspired topoisomerase-based strategies for polynucleotide  
10 synthesis in which DNA oligonucleotides containing CCCTT cleavage sites serve as  
activated linkers for the joining of other DNA molecules with compatible termini (S.  
5 Shuman, *J. Biol. Chem.* 269:32678-32684, 1994). The use of vaccinia topoisomerase  
type I for cloning generally is described in detail in U.S. Patent No. 5,766,891, which  
15 is incorporated by reference herein in its entirety.

Bivalent strand transfer also results in circularization of the acceptor vector  
DNA by placing the topoisomerase cleavage sites on the transfer sequence (a  
20 synthetic bivalent substrate) and cloning the cleaved DNA into the donor vector. This  
strategy is well-suited to the cloning of DNA fragments amplified by PCR. To clone  
PCR products using vaccinia topoisomerase, it is preferred to include a 10 nucleotide  
25 sequence -5'-XXXXAAGGGC- (SEQ ID NO:2) at the 5' end of the two primers used  
for amplification. The 5'-XXXX segment can correspond to any 4-base overhang that  
15 is compatible with the restriction site into which the PCR product will ultimately be  
cloned. The amplification procedure will generate duplex molecules containing the  
sequence 5'-GCCCTTxxxx-3' (SEQ ID NO:3) at both 3' ends (where xxxx is the  
30 complement of XXXX). Incubation of the PCR product with topoisomerase will  
result in cleavage at both termini and allow the covalently activated PCR fragment to  
35 be ligated into the donor vector DNA. From the donor vector the transfer sequence  
can be simultaneously transferred to one or a number of different acceptor vectors  
engineered to contain functional sequences suitable for accomplishing different types  
40 of cloning procedures. For example, an acceptor vector that is a bacterial expression  
vector generally includes a promoter (such as the lac promoter), the Shine-Dalgarno  
25 sequence (for transcription initiation) and the start codon (AUG). Similarly, a  
eukaryotic expression vector includes, but is not limited to, a heterologous or  
45 homologous promoter for RNA polymerase II, a downstream polyadenylation signal,  
the start codon AUG, and a termination codon for detachment of the ribosome.

The donor complex formed upon cleavage by topoisomerase at a 3' proximal  
50 30 site is extremely stable. The transfer sequence can be transferred nearly quantitatively

5 to an acceptor vector with a complementary site even after many hours of incubation  
of the covalent topo-DNA complex at room temperature. The topo-transfer sequence  
complex can even be denatured with 6 M guanidine HCl and then renatured  
10 spontaneously upon removal of guanidine with complete recovery of strand  
transferase activity. Thus, a topoisomerase-activated vector can be prepared once in  
5 quantity and used as many times as needed for preparation of various types of  
acceptor vectors according to the invention.  
15

In addition, two major families of site-specific recombinases from bacteria and  
unicellular yeast have been described: the integrase family and the resolvase/invertase  
20 family. In these recombinases, strand exchange catalyzed by site specific  
recombinases occurs in two steps of (1) cleavage and (2) rejoining, involving a  
covalent protein-DNA intermediate formed between the recombinase enzyme and the  
DNA strand(s). The nature of the catalytic amino acid residue of the enzyme and the  
25 line of entry of the nucleophile is different for these two recombinase families. For  
cleavage catalyzed by the invertase/resolvase family, the nucleophile hydroxyl is  
15 derived from a serine and the leaving group is the 3'-OH of the deoxyribose. For the  
integrase family, the catalytic residue is a tyrosine and the leaving group is the 5'-OH.  
30 In both recombinase families, the rejoining step is the reverse of the cleavage step.

The recombinase activity of Cre has been studied as a model system for the  
35 integrases. Cre is a 38 kD protein isolated from bacteriophage P1. It catalyzes  
recombination at a 34 base pair stretch of nucleic acids called loxP. The loxP site has  
the sequence 5'-ATAACTTCGTATAGCATACATTATACGAAGTTAT-3' (SEQ ID  
40 NO: 4; spacer region underlined), consisting of two 13 base pair palindromic repeats  
flanking an eight basepair core sequence (Hoess *et al.*, *Proc. Natl. Acad. Sci USA*  
25 79:3398, 1982 and U. S. Patent No. 4,959,217, the disclosure of which is herein  
incorporated by reference in its entirety). The repeat sequences act as Cre binding  
45 sites with the crossover point occurring in the internal spacer core. Each repeat  
appears to bind one protein molecule wherein the DNA substrate (one strand) is  
cleaved and a protein-DNA intermediate is formed having a 3'-phosphotyrosine  
50 linkage between Cre and the cleaved DNA strand. Crystallography and other studies  
55

5 suggest that four proteins and two loxP sites (each on a different DNA molecule) form  
a synapsed structure in which the DNA resembles models of four-way Holliday-  
10 junction intermediates, followed by the exchange of a second set of strands to resolve  
the intermediate into recombinant products (see, Guo, *et al.*, *Nature* 389:40-46, 1997).  
5 The asymmetry of the core region of the loxP recombination sequence is responsible  
for directionality of the recombination reaction. When two loxP sites on the same  
15 DNA molecule are in a directly repeated orientation, Cre excises the DNA between  
these two sites, leaving a single loxP site on the DNA molecule (Abremski *et al.*, *Cell*  
32:1301, 1983). Thus, the repeat sequences act as Cre-specific binding sites with the  
10 recombination crossover point occurring in the core.

The loxP site is so complex in size that it occurs only in the P1 phage genome.  
Therefore, use of the loxP sites in the invention vectors assures that the enzyme will  
25 not cut the transfer sequence within the interior of the sequence unless the transfer  
sequence is from the P1 phage genome. The activity of Cre in a wide variety of  
15 cellular backgrounds, including yeast, shows that Cre does not require host specific  
factors for activity (Sauer *Mol. Cell. Biol.* 7:2087-2096, 1987), plants (Albert *et al.*,  
30 *Plant J.* 7:649-659, 1995; Dale and Ow, *Gene* 21:79-85, 1990; Odell *et al.*, *Mol. Gen.*  
*Genet.* 223:369-378, 1990) and mammals, including both rodent and human cells (van  
Deursen *et al.*, *Proc. Natl. Acad. Sci. USA* 92:7376-7380, 1995; Agah *et al.*, *J. Clin.*  
35 *Invest.* 100:169-179, 1997; Sauer and Henderson, *New Biologist* 2:441-449, 1990).

The Cre protein also recognizes a number of variant or mutant lox sites  
(variant relative to the loxP sequence), including the loxB, loxL and loxR sites, which  
40 are found in the *E. coli* chromosome. Other variant lox sites include loxP511  
(5'-ATAACTTCGTATAGTATACATTATACGAAGTTAT-3' (SEQ ID NO:5;  
25 spacer region underlined); loxC2  
(5'-ACAACTTCGTATAATGTATGCTATACGAAGTTAT-3' (SEQ ID NO:6;  
45 spacer region underlined; U.S. Patent No. 4,959,217). Additional variants of the loxP  
site can be prepared by those of skill in the art and will generally have no more than a  
50 total of one to three point mutations in the two repeats that comprise the site-specific  
30 recombination sequence. Cre catalyzes the cleavage of the lox site within the spacer

5 region and creates a six base-pair staggered cut. The two 13 bp inverted repeat  
domains of the lox site represent binding sites for the Cre protein. The two lox sites  
may differ so long as Cre is able to recognize both lox sites. However, if two lox sites  
10 differ in their spacer regions in such a manner that the overhanging ends of the  
cleaved DNA cannot reanneal with one another, Cre cannot efficiently catalyze a  
5 recombination event using the two different lox sites. The efficiency of the  
recombination event will depend on the degree and the location of the variations in the  
binding sites. For example, the loxC2 site can be efficiently recombined with the  
15 loxP site because the two lox sites differ by a single nucleotide in the leftbinding site.  
Thus, when Cre is the site specific recombinase used in the practice of the invention  
20 methods, the site-specific recombination sequence is a loxP site, or a variant thereof  
recognized by the Cre enzyme.

25 A recombinase of the integrase family with similar function is Flp, a  
recombinase identified in strains of *Saccharomyces cerevisiae* that contain 2 $\mu$ -circle  
15 DNA. Flp recognizes a DNA sequence consisting of two 13 basepair inverted repeats  
flanking an 8 basepair core sequence  
30 (5'-GAAGTTCCTATTCTCTAGAAAGTATAGGAAGTTC-3' (SEQ ID NO: 7);  
spacer underlined) called *FRT* (Flp Recombination Target site). A third repeat  
follows at the 3' end in the natural sequence, but does not appear to be required for  
35 recombination activity. The Flp gene has been cloned and expressed in *E. coli* and in  
mammalian cells (PCT International Patent Application PCT/US92/01899,  
Publication No: WO 92/15694, the disclosure of which is herein incorporated by  
reference) and has been purified (Meyer-Lean *et al.*, *Nucleic Acids Res.* 15:6469,  
40 1987; Babincau *et al.*, *J. Biol. Chem.* 260:12313, 1985; Gronostajski and Sadowski, *J.*  
25 *Biol. Chem.* 260:12328, 1985).

45 Like Cre, Flp is functional in a wide variety of systems including bacteria  
(Huang *et al.*, *J. Bacteriology* 179:6076-6083, 1997), insects (Golic and Lindquist,  
*Cell* 52:499-509, 1989; Golic and Golic, *Genetics* 144:1693-1711, 1996), plants  
(Lyznik *et al.*, *Nucleic Acids Res* 21:969-975, 1993) and mammals (U. S. Patent Nos.  
50  
55

5,677,177 and 5,654,182), which shows the FIP does not require host specific factors for operability.

Unlike the integrases, each member of the resolvase subfamily of recombinase enzymes contains an N-terminal catalytic domain having a high degree (>35%) of sequence homology among the subfamily members (Crellin and Rood, *J. Bacteriology* 179(16):5148-5156, 1997; Christiansen *et al.*, *J. Bacteriology* 178(17):5164-5173, 1996). Despite this, like the integrases, many of the resolvases do not require host specific accessory factors (Thorpe and Smith, *PNAS USA* 95:5505-5510, 1998).

Other site-specific recombinases suitable for use in the system and methods of the present invention include RecA (Ferrin *et al.*, *PNAS USA* 95:2156-57, 1998), HK022 integrase, lambda integrase (with or without Xis), which recognizes Att sites (Weisberg *et al.*, In: *Lambda II*, Hendrix *et al.*, Eds., Cold Spring Harbor Press, Cold Spring Harbor, NY, 1983), and the like.

The process of strand exchange used by the resolvases is somewhat different than the process used by the integrases. The resolvases usually make cuts close to the center of the crossover site, and the top and bottom strand cuts are often staggered by 2 basepairs, leaving recessed 5' ends. A protein-DNA linkage is formed between phosphodiester from the 5' DNA end and a conserved serine residue close to the amino terminus of the recombinase. Like the invertases, two proteins units are bound at each crossover site, however, no equivalent to the Holliday-junction intermediate is formed (see Stark *et al.*, *Trends in Genetics* 8(12):432-439, 1992, incorporated by reference herein).

The nucleic acid sequences recognized as recombination sites by members of the resolvase family differ in several ways from the integrases. The sites used for recognition and recombination of the phage and bacterial DNAs (the native host system) are generally non-identical, although they typically have a common core region of nucleic acids. The bacterial sequence is generally called the AttB sequence (bacterial attachment) and the phage sequence is called the AttP sequence (phage

5 attachment). Because AttB and AttP are somewhat different sequences,  
10 recombination will result in a stretch of nucleic acids (called AttL or AttR for left and  
right) that is neither an AttB sequence nor an AttP sequence, and is probably  
unrecognizable as a recombination site to the relevant enzyme, thus reducing the  
5 possibility that the enzyme will catalyze a second recombination reaction that would  
reverse the first.

15 The individual resolvases and the nucleic acid sequences that they recognize  
have been less well characterized than Cre and Flp, although most of the core  
sequences have been identified. The core sequences of some of the resolvases useful  
20 in the practice of the invention include TP901-1 - 5'-TTCAAT(T/C)AAGGTAA  
(SEQ ID NO: 8); TnpX - 5'-GCCCNGA(G/A)GG (SEQ ID NO: 9), R4 - 5'-  
GAAGCAGTGGTA (SEQ ID NO: 10), and  $\phi$ C31 - 5'-TTG (SEQ ID NO: 11) (see  
25 Rausch and Lehmann, *NAR* 19:5187-5189, 1991; Shirai *et al.*, *J. Bacteriology*  
173(13):4237-4239, 1991; Crellin and Rood, *J Bacteriology* 179:5148-5156, 1997;  
15 Christiansen *et al.*, *J Bacteriology* 176:1069-1076, 1994, all of which are incorporated  
by reference herein.)

30 In general, Site-specific recombination sequences of the invention vary in  
length, although they are generally less than 50 nucleotides. Particularly suitable site-  
specific recombination sequences include the recognition sequences for vaccinia  
35 topoisomerase I (5'-(C/T)CCTT↓, SEQ ID NO: 1), Cre (5'-ATAACTTCGTATA  
GCATACAT TATACGAAGTTAT-, SEQ ID NO: 4), Flp (5'-GAAGTTCCTATAC  
TTCTAGAA GAATAGGAAC TTC, SEQ ID NO: 7), lambda integrase  
40 (5'-CAAGTT, SEQ ID NO: 12), HK022 integrase (5'-AACCTT, SEQ ID NO: 13),  
and the like. The present invention is illustrated, but not limited by the use of vectors  
25 containing topoisomerase I sites.

45 Any nucleic acid sequence is suitable as a transfer sequence as long as there is  
a desire for the sequence to be moved from one vector to another. The transfer  
sequence may, for example, encode a protein, peptide or functional RNA (such as  
50 antisense sequences, hammerhead ribozymes, and the like). A transfer sequence



5 encoding a protein or peptide may be either a gene sequence or a coding sequence. As  
used herein, a "gene sequence" is the entire nucleic acid sequence that is necessary for  
10 the synthesis of a functional polypeptide or RNA molecule; whereas a "coding  
sequence" is limited to the nucleic acids encoding the amino acid sequence of a  
5 protein.

15 The transfer sequence may also be a sequence whose function, if any, is not  
yet known, such as an expressed sequence tag (EST) fragment. Such sequences can  
be used as diagnostic probes, or as aids in the identification and cloning of a larger  
sequence containing the EST fragment.

20 10 The vectors employed in the practice of the invention contain one or more  
nucleic acid sequences in addition to the site-specific recombination sequences, and  
transfer sequence in the case of a donor vector. The additional nucleic acid sequences  
25 will generally have some function in the replication or integrity of the vector, in the  
expression of a protein, in the modification of an expressed protein, and the like.

15 15 Particularly useful nucleic acid sequences include promoter-enhancer sequences,  
selection marker sequences, origins of replication, inducible element sequences,  
30 fusion protein producing sequences, for example, localization signal sequences,  
epitope tags, proteolytic cleavage recognition sequences, polypeptides that facilitate  
purification, and the like.

35 20 Promoter-enhancer sequences are DNA sequences to which RNA polymerase  
binds and initiates transcription. The promoter determines the polarity of the  
transcript by specifying which strand will be transcribed. Bacterial promoters consist  
40 of consensus sequences, -35 and -10 nucleotides relative to the transcriptional start,  
which are bound by a specific sigma factor and RNA polymerase. Eukaryotic  
25 promoters are more complex. Most promoters utilized in vectors are transcribed by  
45 RNA polymerase II. General transcription factors (GTFs) first bind specific  
sequences near the start and then recruit the binding of RNA polymerase II. In  
addition to these minimal promoter elements, small sequence elements are recognized  
50 specifically by modular DNA-binding/trans-activating proteins (e.g. AP-1, SP-1) that

5 regulate the activity of a given promoter. Viral promoters serve the same function as  
bacterial or eukaryotic promoters and either provide a specific RNA polymerase in  
trans (bacteriophage T7) or recruit cellular factors and RNA polymerase (SV40, RSV,  
10 CMV). Viral promoters may be preferred as they are generally particularly strong  
5 promoters.

15 Promoters may be, furthermore, either constitutive or regulatable (i.e.,  
inducible or derepressible). Inducible elements are DNA sequence elements which  
act in conjunction with promoters and bind either repressors (e.g. lacO/LAC Iq  
repressor system in *E. coli*) or inducers (e.g. gal1/GAL4 inducer system in yeast). In  
20 either case, transcription is virtually "shut off" until the promoter is derepressed or  
induced, at which point transcription is "turned-on".

25 Examples of constitutive promoters include the int promoter of bacteriophage  
 $\lambda$ , the bla promoter of the  $\beta$ -lactamase gene sequence of pBR322, the CAT promoter  
of the chloramphenicol acetyl transferase gene sequence of pPR325, and the like.  
15 Examples of inducible prokaryotic promoters include the major right and left  
promoters of bacteriophage ( $P_L$  and  $P_R$ ), the trp, reca, lacZ, LacI, AraC and gal  
promoters of *E. coli*, the  $\alpha$ -amylase (Ulmanen *et al.*, *J. Bacteriol.* 162:176-182, 1985)  
30 and the sigma-28-specific promoters of *B. subtilis* (Gilman *et al.*, *Gene Sequence*  
32:11-20, 1984), the promoters of the bacteriophages of *Bacillus* (Gryczan, In: *The*  
35 *Molecular Biology of the Bacilli*, Academic Press, Inc., NY, 1982), *Streptomyces*  
20 promoters (Ward *et al.*, *Mol. Gen. Genet.* 203:468-478, 1986), *Pichia* promoters  
(U.S. Patent Nos. 4,855,231 and 4,808,537), and the like. Exemplary prokaryotic  
promoters are reviewed by Glick (*J. Ind. Microbiol.* 1:277-282, 1987); Cenatiempo  
40 (*Biochimie* 68:505-516, 1986); and Gottesman (*Ann. Rev. Genet.* 18:415-442, 1984).

25 Preferred eukaryotic promoters include, for example, the promoter of the  
mouse metallothionein I gene sequence (Harner *et al.*, *J. Mol. Appl. Gen.* 1:273-288,  
45 1982); the TK promoter of Herpes virus (McKnight, *Cell* 31:355-365, 1982); the  
SV40 early promoter (Benoist *et al.*, *Nature (London)* 290:304-310, 1981); the yeast  
50 *gal1* gene sequence promoter (Johnston *et al.*, *Proc. Natl. Acad. Sci. (USA)* 79:6971-

5 6975, 1982); Silver *et al.*, *Proc. Natl. Acad. Sci. (USA)* **81**:5951-5955, 1984), the  
CMV promoter, the EF-1 promoter, Ecdysone-responsive promoter(s), tetracycline-  
10 responsive promoter, and the like.

15 Selection marker sequences are valuable elements in expression vectors as  
5 they provide a means to select for growth only those cells which have been  
successfully transformed with a vector containing the selection marker sequence and  
express the marker. Such markers are of two types: drug resistance and auxotrophic.  
A drug resistance marker enables cells to detoxify an exogenously added drug that  
would otherwise kill the cell. Auxotrophic markers allow cells to synthesize an  
20 10 essential component (usually an amino acid) while grown in media which lacks that  
essential component.

25 Common selectable marker gene sequences include those for resistance to  
antibiotics such as ampicillin, tetracycline, kanamycin, bleomycin, streptomycin,  
hygromycin, neomycin, Zeocin™, and the like. Selectable auxotrophic gene  
15 sequences include, for example, *hisD*, which allows growth in histidine free media in  
the presence of histidinol.

30 A further element useful in a vector is an origin of replication sequence.  
Replication origins are unique DNA segments that contain multiple short repeated  
35 sequences that are recognized by multimeric origin-binding proteins and which play a  
20 key role in assembling DNA replication enzymes at the origin site. Suitable origins of  
replication for use in expression vectors employed herein include *E. coli oriC*, *colE1*  
plasmid origin, 2 $\mu$  and ARS (both useful in yeast systems), *sfi*, SV40 EBV oriP  
40 (useful in mammalian systems), and the like.

45 Fusion protein producing sequences may be included in a vector employed in  
25 the present invention. When two protein-coding sequences not normally associated  
with each other in nature are in the same reading frame the resulting expressed protein  
is called a "fusion protein" as two distinct proteins and/or fragments have been  
"fused" together. Fusion proteins have a wide variety of uses. For example, two  
50 functional enzymes can be fused to produce a single protein with multiple enzymatic

5 activities or short peptide sequences can be fused to a larger protein and serve as aids  
in purification or as means of identifying expressed protein by serving as epitopes  
detectable by specific antibodies. Thus examples of fusion protein producing  
10 sequences useful in the vectors of the invention include epitope-tag encoding  
5 sequences, affinity purification-tag encoding sequences, functional protein encoding  
sequences, and the like.

15 Epitope tags are short peptide sequences that are recognized by epitope  
specific antibodies. A fusion protein comprising a recombinant protein and an epitope  
tag can be simply and easily purified using an antibody bound to a chromatography  
20 resin. The presence of the epitope tag furthermore allows the recombinant protein to  
be detected in subsequent assays, such as Western blots, without having to produce an  
antibody specific for the recombinant protein itself. Examples of commonly used  
25 epitope tags include V5, glutathione-S-transferase (GST), hemagglutinin (HA), the  
peptide Phe-His-His-Thr-Thr, chitin binding domain, and the like.

15 Affinity purification tags are generally peptide sequences that can interact with  
a binding partner immobilized on a solid support. Preferably, the recombination event  
30 in the invention method places the transfer sequence in frame with the sequence  
encoding the affinity domain, so that the affinity purification tag and the expression  
product of the transfer sequence is expressed as a fusion protein when the sequence is  
35 expressed. DNA sequences encoding multiple consecutive single amino acids, such  
as histidine, when fused to the expressed protein, may be used for one-step  
20 purification of the recombinant protein by high affinity binding to a resin column,  
such as nickel sepharose. An endopeptidase recognition sequence can be engineered  
40 between the polyamino acid tag and the protein of interest to allow subsequent  
25 removal of the leader peptide by digestion with enterokinase, and other proteases.  
Sequences encoding peptides, such as the chitin binding domain (which binds to  
chitin), glutathione-S-transferase (which binds to glutathione), biotin (which binds to  
45 avidin and strepavidin), and the like, can also be used for facilitating purification of  
the protein of interest. The affinity purification tag can be separated from the protein  
50

of interest by methods well known in the art, including the use of inteins (protein self-splicing elements, Chong *et al.*, *Gene* **192**:271-281, 1997).

The use of the term "functional protein encoding sequence", as used herein, indicates that the fusion protein producing element of a vector encodes a protein or peptide having a particular activity, such as an enzymatic activity, a binding activity, and the like. For example, a functional protein encoding sequence may encode a kinase catalytic domain (Hanks and Hunter, *FASEB J* **9**:576-595, 1995), producing a fusion protein that can enzymatically add phosphate moieties to particular amino acids, or may encode a Src Homology 2 (SH2) domain (Sadowski, *et al.*, *Mol. Cell. Bio.* **6**:4396, 1986; Mayer and Baltimore, *Trends Cell. Biol.* **3**:8, 1993), producing a fusion protein that specifically binds to phosphorylated tyrosines.

The foregoing elements can be combined to produce vectors suitable for use in the methods of the invention. Those of skill in the art would be able to select and combine the elements suitable for use in any particular system.

Suitable prokaryotic vectors include plasmids such as those capable of replication in *E. coli* (for example, pBR322, ColEI, pSC101, PACYC 184, itVX, pRSET, pBAD (Invitrogen, Carlsbad, CA), and the like). Such plasmids are disclosed by Sambrook (cf. *Molecular Cloning: A Laboratory Manual*, second edition, edited by Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989). *Bacillus* plasmids include pCI94, pC221, pT127, and the like, and are disclosed by Gryczan (In: *The Molecular Biology of the Bacilli*, *supra*, pp. 307-329). Suitable *Streptomyces* plasmids include pIJ101 (Kendall *et al.*, *J. Bacteriol.* **169**:4177-4183, 1987), and streptomyces bacteriophages such as  $\phi$ C31 (Chater *et al.*, In: *Sixth International Symposium on Actinomycetales Biology*, Akademiai Kiado, Budapest, Hungary, pp. 45-54, 1986). *Pseudomonas* plasmids are reviewed by John *et al.* (*Rev. Infect. Dis.* **8**:693-704, 1986), and Izaki (*Jpn. J. Bacteriol.* **33**:729-742, 1978).

Suitable eukaryotic plasmids include, for example, BPV, EBV, vaccinia, SV40, 2-micron circle, pcDNA3.1, pcDNA3.1/GS, pYES2/GS, pMT, p IND, pIND(Sp1), pVgRXR (Invitrogen), and the like, or their derivatives. Such plasmids

are well known in the art (Botstein *et al.*, *Miami Wntr. Symp.* 12:265-274, 1982; Broach, In: *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY pp. 445-470, 1981; Broach, *Cell* 28:203-204, 1982; Dilon *et al.*, *J. Clin. Hematol. Oncol.* 10:39-48, 1980; Maniatis, In: *Cell Biology: A Comprehensive Treatise*, Vol. 3, Gene Sequence Expression, Academic Press, NY, pp. 563-608, 1980.

A further embodiment of the invention comprises a method of rapidly subcloning a nucleic acid sequence. The invention method comprises contacting a site-specific recombinase and a cell-free solution comprising a donor vector comprising a transfer sequence flanked by a site-specific recombination sequence recognized by the recombinase, and an acceptor vector comprising at least one site-specific recombination sequence recognized by the recombinase, under conditions suitable to promote the transfer of the transfer sequence from the donor vector to the acceptor vector. The invention method employs vectors and recombinases as described above. Means of identifying conditions for the transfer of a transfer sequence from a donor vector to an acceptor vector can readily be determined by those of skill in the art. Suitable conditions include those described in Nunes-Düby *et al.*, *EMBO J.* 13(18):4421-4430, 1994; Senecoff *et al.*, *PNAS USA* 82:7270-7274, 1985; Shaikh and Sadowski, *J. Biol. Chem.* 272(9):5695-5702, 1997; and Peterson and Shuman, *J. Biol. Chem.* 272(7):3891-3896, 1997, all of which are incorporated by reference herein, and are described in detail in the Examples set out below.

For example, the invention method can be used to perform subcloning (transfer of a DNA or RNA sequence from one vector to another) without PCR amplification using topoisomerase, as described in Examples 1A-C below. In this embodiment of the invention, donor vector is constructed as shown in Figure 1 with recognition sites for vaccinia topoisomerase I flanking the insertion point for the transfer sequence. The two recognition sites are juxtaposed on opposite strands of the DNA and are generally separated by about four spacer nucleotides to provide overhang. The spacer nucleotides have identical sequences on either side of the insertion point for the gene of interest. One or more vectors is prepared as a linear,

5 double-stranded molecule with single strand overhangs that are compatible with the  
spacer sequences that flank the gene or gene fragment of interest on the donor vector,  
10 as shown in Figure 1B. In addition, the linear acceptor vector DNA has 5'-hydroxyl  
groups at each end. A marker gene sequence and additional sequences are included in  
5 the acceptor vector as known in the art depending upon the particular attribute of the  
vector desired. Multiple acceptor vectors useful for different cloning tasks can be  
15 simultaneously prepared by including in each those attributes suitable to the task for  
which the vector would be used.

20 The donor vector(s) are treated with topoisomerase I for five minutes at room  
10 temperature. The enzyme generates nicks at each topoisomerase recognition site,  
creating double strand breaks at the sites that flank the inserted gene or gene fragment  
of interest and releasing the transfer DNA fragment. Topoisomerase I is covalently  
25 attached at each end of the freed DNA fragment, which also has overhangs  
complementary to the spacer nucleotides. The topoisomerase treated vector is  
15 combined with the linearized acceptor vector in a suitable medium. The compatible  
ends of each vector corresponding to the spacer sequence brings the two DNA  
30 fragments together and allows the topoisomerase I to ligate the spacer sequences  
together in an ATP independent ligation. The recombinant vector formed, shown in  
Figure 1C, contains the gene or gene fragment of interest and can be identified  
35 20 following transformation of the vector into competent E. coli by expression of the  
marker gene. When Cre or Flp is used as the site specific recombinase, the donor and  
acceptor vectors are prepared as described in Example 1 except that the recognition  
40 sites appropriate to the recombinase of choice flank the insertion point for the gene of  
interest.

25 In another embodiment of the invention method, a gene or gene fragment  
donor clones are created by PCR amplification cloning using primers designed for the  
45 specific fragment of interest. A donor vector is not needed. The gene or gene  
fragment of interest is generated repeatedly from the donor clone for insertion into any  
or all of the acceptor vectors for a wide variety of research or production applications.  
50 30 No subcloning is required in this technique to move the gene of interest from one

vector into another. The gene or gene fragment is simply copied from a donor clone, and the copies are inserted into a "copy ready vector" using the following procedure.

In this procedure, the exact sequence of the open reading frame, if any, and of native features of the gene to be transferred should be noted if the gene is to be expressed as a fusion protein from one or more of the acceptor vectors. For example, signal sequences for intracellular organelle targeting, secretion, glycosylation, etc. are identified in the transfer sequence to determine that the gene of interest is in reading frame with any signal sequence or genes encoding a tag, and the like, in the acceptor vector.

Oligonucleotides are designed for PCR amplification of the exact DNA sequence to be transferred to the acceptor vector(s) using one or more methods well known in the art. For example, to transfer a complete open reading frame, the sequence of one oligonucleotide would have the translation initiation codon at its 5'-end and the sequence of the other oligonucleotide would have the translation initiation codon at its 5'-end. The sequence of the other oligonucleotide would have the complement of the translation termination codon at its 3'-end. Acceptor vectors are prepared as described in Example 1, such as an acceptor vector including DNA sequences appropriate for the expression or analysis of the protein encoded by the gene of interest.

The gene sequence of interest is amplified from the donor clone using the PCR primers prepared as above-described, with cycling parameters selected as suitable for the primer and the template. A 7 to 30 minute extension at 72° C is optionally included to ensure that all amplified products are full length and 3' adenylated. The amplified DNA fragment is ligated into the acceptor vector(s). In general, 0.5 to 2 µl of the PCR product (10 ng/µl) with an average insert length of 400 to 100 bp gives a proper insert:vector ratio. Therefore the PCR product is ligated into the acceptor vector by placing 0.5 to 2 µl of PCR product reaction in sterile water to provide a final volume of 4 µl. To this mixture is added 1 µl of the acceptor vector to obtain a final volume of 5 µl., mixing gently and incubating for 5 minutes at room temperature



(25° C), then centrifuging briefly and placing the tube on ice. Competent cells, such as *E. coli*, are then immediately transformed with the acceptor vector(s).

In yet another embodiment of the invention method, gene or gene fragment clones are created by PCR amplification using primers designed specifically, or non-specifically, for the fragment, but which also contain sequences that, when the amplified gene fragment is inserted into an invention donor vector, will allow use of a universal donor vector primer set to create copies of the gene or gene fragment for insertion into one or more specialty application acceptor vectors using the following procedure. If a collection of genes are to be transferred, each gene of interest should be available on a donor plasmid vector and flanked by short sequences that are common to all donor plasmids in the collection. Oligonucleotides for PCR amplification of the gene(s) are synthesized based on the short sequence that flanks each of the transfer sequences in the donor vectors.

An invention acceptor vector containing a recombinase recognition site appropriate for the expression or analysis of the gene of interest is selected. For example, the acceptor vector containing a topoisomerase I recognition site, a strong mammalian promoter, and the coding sequence for an epitope tag would be appropriate for production and analysis of the protein of interest, such as the TOPO Cloning™ vector (Invitrogen, Carlsbad, CA). The transfer sequence(s) of interest are amplified from the donor vector using the PCR primers with cycling parameters suitable for the particular primers and template. It may be necessary to include a 7 to 30 minute extension of 72°C to ensure that all amplified products are full length and 3' adenylated. The amplified DNA fragments are individually transferred into acceptor vectors) using the insert:vector ratio and conditions described above.

In a presently preferred embodiment of the invention method, the PCR primers add the following sequences at the 5' end to add topoisomerase I recognition sites to the ends of the amplified PCR product:

Forward Primer      5'-AAGGG (SEQ ID NO:14)

## Reverse Primer 5'-CCCTT (SEQ ID NO:1)

The acceptor vector is prepared as a linear molecule with single 3'-T overhangs and 5'-hydroxyl groups. After amplification by PCR, the PCR product is treated with topoisomerase I so that the enzyme becomes covalently bound to each end of the amplified PCR product. Then the covalently bound PCR product is introduced into the acceptor vector(s) as described above.

In another embodiment, the invention provides kits comprising one or more containers or vials containing components for carrying out the methods of the present invention. For instance, such a kit can comprise a suitable reaction solution, recombinase and cells. Also included in the kit are one or more vectors, *e.g.*, vectors for expression in mammalian, bacterial, yeast and insect cells. In a preferred embodiment, the kit will comprise a reaction solution of 50 mM Tris HCl pH 7.5, one or more of the invention vectors that have vaccinia DNA topoisomerase covalently bound thereto, and instructions for their use as described herein.

In one embodiment the invention kit comprises at least one donor vector comprising at least one site specific recombination sequence, a transfer sequence, and a first selectable marker, and at least one acceptor vector comprising at least one site specific recombination sequence, a lethal gene and a second selectable marker. For example, as illustrated in Figure 2, the donor vector in the kit can contain a selectable marker gene other than Zeocin, an origin of replication sequence ("ori"), and a transfer sequence ("gene of interest") flanked by lox P recognition sites. The acceptor vector ("pAcceptor") then contains a gene encoding resistance to the antibiotic Zeocin™ ("Zeo"), an origin of replication sequence ("ori"), and a gene encoding ccdB, a lethal compound, flanked by loxP sites. When the donor and acceptor vectors are combined in a reaction mixture in the presence of the recombinase Cre, a new recombinant vector ("pRecombinant") is created, which recombinant vector contains the transfer sequence and a gene encoding Zeo. Cells transformed with the reaction mixture will grow in the presence of the antibiotic Zeocin™ only if the recombination event has successfully occurred.

5

10

Vaccinia DNA topoisomerase can be prepared for expression in *E. coli* and purified as described in S. Shuman *et al.*, *J. Biol. Chem.* **263**:16401-16407, 1988.

15

The invention will now be described in greater detail by reference to the following non-limiting Examples.

#### EXAMPLE 1

##### Subcloning without PCR Amplification using topoisomerase

20

25

Donor vectors can be constructed such that recognition sites for topoisomerase, or other ATP independent enzymes, flank the transfer sequence. In the presence of acceptor vector and topoisomerase, or other ATP independent enzyme, the transfer sequence is occasionally subcloned from the donor vector to the acceptor vector in an ATP independent event.

30

35

40

45

50

55

A linear activated vector containing vaccinia topoisomerase recognition sites (e.g., pCR2.1-TOPO (Invitrogen)) is prepared to receive the transfer sequence. The transfer sequence is amplified from a DNA template of choice. The DNA template may be genomic DNA, plasmid DNA, cosmid DNA or any other shuttle construct. Isolation methods are available in the public domain (Ausubel *et al.*, Section 2.14). Specific oligos (primers) for PCR corresponding to the exact sequence of the transfer DNA are synthesized according to published protocols (Ausubel *et al.*, Section 2.11). Both primers contain 7-9 additional bases on the 5' ends including the complement to the vaccinia topoisomerase I recognition site (5'-AAGGG 3') and an additional 2-4 bases which will serve as the 5' overhangs during subcloning with topoisomerase (5'-CGAAGGG . . . 3', SEQ ID NO:15). PCR amplification is performed utilizing methods optimized for the template and primers (Ausubel *et al.*, Section 15.1) with a DNA polymerase containing terminal transferase activity, such as Taq (Boehringer Mannheim, Indianapolis, IN).

5  
10  
15  
Approximately 20ng of PCR product is combined with 1µl of the prepared activated vector in a total volume of 5µl. The reaction is incubated at 25°C for 5 min, placed on ice and 1µl is transformed into competent *E. coli* using either chemical transformation or electroporation techniques (Ausubel *et al.*, Section 1.8). Transformed cells are  
5 plated on appropriate antibiotic selection plates and grown at 37°C for 12-18 hours. Resulting colonies are screened by miniprep and restriction digest (Ausubel *et al.*, Sections 1.6 and 3.1) to identify clones containing transfer sequence.

20  
25  
15  
Positive clones will contain the transfer sequence flanked on each side by 2 tandem topoisomerase recognition sites on complementary strands separated by 2-4 bases (for example, a direct repeat of 5'-CCCTTGCAAGGG (SEQ ID NO:16) with an  
10 intervening transfer sequence). A positive clone is propagated in *E. coli* and the plasmid DNA is purified as described above. The plasmid DNA is resuspended in TE Buffer, pH 8 (10mM Tris, 1mM EDTA) at a concentration of 10 ng/µl. This vector  
25 will serve as the donor for subcloning in an ATP independent reaction using  
15 topoisomerase.

#### 30 B. Preparation of the Acceptor Vector

35  
40  
45  
20  
Preparation of linear, dephosphorylated vector: Supercoiled plasmid DNA to be used for construction of the acceptor vector is propagated and purified as described above. The plasmid chosen to be the acceptor vector must have a different  
25 *E. coli* antibiotic selection marker from the donor vector, for example Zeocin™. Plasmid DNA is digested with a restriction enzyme that is unique within the vector and will leave the desired 2-4 base 5' overhangs (e.g., digestion with *BstB I* will leave  
40 2 base 5' overhangs). It is possible to digest with two different enzymes for directional cloning, however the forward and reverse PCR primers used to create the  
25 donor vector must be designed to generate the necessary complementary overhangs.

50  
55  
Plasmid DNA (30µg) is digested with 120 units of *BstB I* (New England BioLabs, Beverly, MA) for 2 hours under conditions specified by the supplier, extracted with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1), ethanol precipitated, and washed with 500µl of 80% ethanol (Ausubel *et al.*, Section

5 2.1). The DNA ends are dephosphorylated by treating with calf intestinal alkaline  
phosphatase (CIP; New England BioLabs, Beverly, MA) according to protocol  
10 specified by the supplier, extracted with phenol/chloroform/isoamyl alcohol (25:24:1),  
ethanol precipitated, and washed with 80% ethanol (Ausubel *et al.* Section 2.1). The  
5 DNA is resuspended in 1000 µl of TE buffer, pH 8.

### 15 C. Subcloning with Topoisomerase

Cell-free subcloning and selection: 10ng of prepared donor vector, 30ng of  
prepared acceptor vector and 1 µg of purified topoisomerase are combined in a total  
20 volume of 5 µl, and incubated for 5 min. at 25°C to allow transfer of the desired  
sequence from the donor vector to the acceptor vector in an ATP independent  
10 reaction. The reaction mixture is placed on ice and 1 µl is transformed into competent  
*E. coli* using either chemical transformation or electroporation techniques (Ausubel *et*  
25 *al.*, Section 1.8). Clones containing acceptor vector plus transfer sequence are  
selected by plating on antibiotic media requiring a resistance marker specific to the  
15 acceptor vector (e.g., Zeocin™). Plates are incubated at 37°C for 12-18 hours.  
Resulting colonies are screened by miniprep and restriction digest (Ausubel *et al.*,  
30 Sections 1.6 and 3.1) to identify clones containing the desired transfer sequence  
subcloned into the acceptor vector.

## 35 EXAMPLE 2

### 20 Protocol 2: Subcloning without PCR Amplification Using Site-Specific Recombinases

#### 40 A. Preparation of Donor Vector

Construct Design: A donor vector is constructed so that a transfer sequence  
45 and a unique bacterial selection marker (e.g. Zeocin™, Invitrogen Corp., Carlsbad,  
25 CA) are flanked by tandemly repeated recombinase recognition sites (for example  
*loxP* or *FRT*). The donor vector construct containing recombinase recognition sites is  
50 built using standard molecular biology techniques of PCR and subcloning (Ausubel *et*

5 *al.*, Sections 3.16 and 3.17). The desired transfer sequence may be subcloned into the donor vector using either standard PCR/restriction digest and ligation techniques  
10 (Ausubel *et al.*, Sections 3.16 and 3.17) or by topoisomerase mediated cloning of PCR products as described in Examples 3 and 5 hereafter.

5 **Donor Vector Preparation:** The donor plasmid DNA is propagated in *E. coli* (see Example 1: Section A) and purified from 100 ml of a saturated culture according to protocols specified for the SNAP<sup>TM</sup> Midiprep Kit (Invitrogen, Carlsbad, CA). The  
15 plasmid DNA is resuspended in TE Buffer, pH 8 (10mM Tris, 1mM EDTA) at a concentration of 0.5µg/µl.  
20

#### 25 **B. Preparation of the Acceptor Vector**

**Construct Design:** The acceptor vector contains a single recombination  
30 recognition site in the desired cloning region that is identical to the two sites on the donor vector. It also contains a bacterial selection marker that differs from that of the  
15 donor vector (e.g., Ampicillin) to allow for selection of acceptor vector clones. The acceptor vector is built using standard molecular biology techniques of PCR and subcloning (Ausubel *et al.*, Sections 3.16 and 3.17).  
35

**Acceptor Vector Preparation:** The acceptor plasmid DNA is propagated in  
20 *E. coli* (Example 1: Section A above) and purified from 100ml of a saturated culture according to protocols specified for the SNAP<sup>TM</sup> Midiprep Kit (Invitrogen, Carlsbad, CA). The plasmid DNA is resuspended in TE Buffer pH 8 at a concentration of 0.5µg/µl.  
40

#### 45 **C. Subcloning with a Site-Specific Recombinase**

25 **Recombinase (Cre) Reaction:** A combination of 0.25µg of donor vector, 0.75µg of acceptor vector, 6µl of 10X Cre Buffer (50mM Tris-HCl, pH 7.5, 33mM  
50

NaCl, 10mM MgCl<sub>2</sub>, 100µg/ml BSA) and 2 units of Cre Recombinase (Novagen, Madison, WI) is prepared in a 60µl total volume and incubated at 37°C for 15 min. Competent *E. coli* are transformed with 2µl of the combination using either chemical transformation or electroporation techniques (Ausubel *et al.*, Section 1.8). Based on incompatibility of different vectors containing the same origin of replication within a single cell (*Molecular Cloning, A Laboratory Manual*, Second Edition, Ed. Sambrook *et al.*, Cold Spring Harbor Laboratory Press, New York, 1989, p. 1.4), clones containing acceptor vector plus transfer sequence are selected by plating on antibiotic media requiring resistance markers specific to both the acceptor vector and the donor vector region that is subcloned (e.g., Ampicillin and Zeocin™). Plates are incubated at 37°C for 12-18 hours. The resulting colonies are screened by miniprep and restriction digest (Ausubel *et al.*, Sections 1.6 and 3.1) to identify clones containing the desired transfer sequences and subcloned into the acceptor vector.

### EXAMPLE 3

#### Cloning PCR amplified DNA with gene specific primers and Cloning vector.

Gene or gene fragment amplimers are created by PCR amplification using primers sequence-specific to the gene or gene of interest. Any region of DNA containing the gene of interest (designated the donor) and primers specific to the gene of interest can be used to generate the amplimer repeatedly for insertion into any or all of the acceptor vectors for a wide variety of research or production applications. No subcloning is required in this technique to transfer the gene or gene fragment of interest into the acceptor vector. The amplimer is simply copied off from the donor and the copies inserted into the acceptor vector using the procedure described below.

5 The DNA template should be available in sufficient quantities (at least 20 ng for plasmids) and the complete sequence of the target open reading frame should be known. DNA template may be genomic DNA, plasmid DNA, cosmid DNA or any  
10 other shuttle construct. Isolation methods are those known in the art, for example, as disclosed in Ausubel *et al.*, Section 2.14.

15 Specific oligonucleotides (primers) for PCR corresponding to the exact DNA sequence to be transferred to the acceptor vector are prepared. For example, to transfer a complete open reading frame, the sequence of the 5' primer would contain the translation initiation codon and flanking sequences of the target sequence. The  
20 10 sequence of the 3' primer would contain the complement of the translation termination sequence of the target. Protocols describing the synthesis of oligonucleotides are available in the public domain (Ausubel *et al.*, Section 2.11).

25 An acceptor vector appropriate for the expression or analysis of the gene or gene fragment of interest is TOPO Cloning™ vector, having the topoisomerase already  
15 associated with the linear plasmid, for example, pCR2.1TOPO™ (Invitrogen,

30 The transfer sequence of interest is obtained from the donor clone in a 50 µl reaction volume using the PCR primers specific to the transfer sequence. Cycling parameters are selected to be appropriate for the primers and template used (Ausubel  
35 *et al.*, Section 15.1). It may be necessary to include a 7 to 30 minute extension at 20 72°C after PCR is complete to ensure that all amplimers are full length and 3' adenylated (Ausubel *et al.*, Section 15.7).

40 The amplimer is cloned into the acceptor vector as follows: For one reaction, 0.5 to 2 µl fresh PCR product is combined with 1 µl of the acceptor vector and sterile water is added to a 5 µl total volume. The mixture is gently  
45 25 stirred and incubated for 5 minutes at room temperature (~25°C) and then competent *E. coli* cells are immediately transformed with the mixture by any known method.



In general, 0.5 to 2  $\mu$ l of a typical PCR reaction (10 ng/ $\mu$ l) with an average amplicon length of 400 to 1000 bp will give the proper insert:vector ratio.

#### EXAMPLE 4

##### Cloning PCR amplified DNA with generic primers and a cloning vector.

Gene or gene fragment amplicons are created by PCR amplification using primers of sequence specific to the donor vector and unrelated to the transfer sequence (generic). Any plasmid containing the transfer sequence (designated the donor plasmid) and primers specific to the donor plasmid can be used to generate the amplicon repeatedly for insertion into any or all of the acceptor vectors for a wide variety of research or production applications. No subcloning is required in this technique to transfer the gene or gene fragment of interest into the acceptor vector. The amplicon is simply copied off from the donor plasmid and the copies inserted into the acceptor vector using the procedure described below.

The donor plasmid should be available in sufficient quantities (at least 20 ng) and the complete sequence of the target open reading frame should be known. Isolation methods are well known in the art (Ausubel *et al.*, Section 2.14).

Specific oligonucleotides (primers) are prepared corresponding to the plasmid DNA sequences flanking the amplicon to be transferred to the acceptor vector. Primers need to be made corresponding to regions of the plasmid immediately upstream and downstream of the amplicon. Protocols describing the synthesis of oligonucleotides are well known in the art (Ausubel *et al.*, Section 2.11).

An acceptor vector appropriate for the expression or analysis of the gene or gene fragment of interest and having the topoisomerase already associated with the linear plasmid is prepared. Such vectors are commercially available as TOPO Cloning™ vector, having the topoisomerase already associated with the linear plasmid, for example, pCR2.1 TOPO™ (Invitrogen, Carlsbad, CA).

5 The transfer sequence is cloned from the donor plasmid in a 50 µl reaction  
volume using the PCR primers specific to the donor plasmid, and utilizing cycling  
10 parameters that are appropriate for the primers and template as described in Example  
3 above. The amplimer is cloned into the acceptor vector as described in Examples 1-  
5 3 above.

### 15 EXAMPLE 5

#### 20 Transferring PCR amplified DNA treated with topoisomerase.

25 A desired transfer sequence is amplified from a donor clone by PCR using  
10 primers specific for the transfer sequence. The inclusion of topoisomerase recognition  
sites at the 5' ends of the PCR primers enables transfer of the amplified sequence to  
an appropriate acceptor vector when treated with topoisomerase.

#### A. PCR Amplified Transfer DNA

30 Preparation of amplified transfer DNA treated with topoisomerase: A  
15 donor clone may be genomic DNA, cDNA, plasmid DNA, cosmid DNA or any other  
shuttle construct. DNA from the donor clone is prepared for use as a template in PCR  
amplification utilizing an appropriate preparation technique (Ausubel et. al., Sections  
35 2.11 and 5.5). The sequence of the transfer DNA is known. DNA PCR primers  
containing the complement of the vaccinia topoisomerase I recognition site (SEQ ID  
20 NO:14 followed by transfer DNA specific sequence are synthesized according to  
known protocols (Ausubel et. al, Section 2.11). The DNA fragment generated using  
40 these primers will contain topoisomerase recognition sites at the 3' ends. An  
additional 2-4 bases may be added at the 5' ends of each primer to create 5' overhangs  
45 in the amplified DNA after treatment with topoisomerase. For example, including  
25 SEQ ID NO:14 in the primer will result in 5' overhangs complementary to those  
generated by digestion with *EcoR I*.

5 The transfer sequence is amplified by PCR following established methods  
(Ausubel et. al, Section 15.1). 200ng of amplification product, 200ng of purified  
10 topoisomerase I and TE buffer, pH 8 (10mM Tris, 1mM EDTA) are combined in a  
total volume of 20 $\mu$ l. The reaction is incubated at 25°C for 5 min and placed on ice.  
5 The topoisomerase will be covalently bound to the 3' ends of the PCR product,  
leaving the desired 5' overhangs.

#### 15 B. Preparation of the Acceptor Vector

Preparation of linear, dephosphorylated vector: Supercoiled plasmid DNA  
20 to be used for construction of the acceptor vector is propagated and purified as  
described previously (Example 1, Section A). The plasmid chosen to be the acceptor  
10 vector should have a different *E. coli* antibiotic selection marker from the donor  
vector. Plasmid DNA is digested with a restriction enzyme that is unique within the  
25 vector and will leave the desired 2-4 base 5' overhangs (e. g. digestion with *EcoR I*  
will leave 4 base 5' overhangs: 5'-AATT...-3'). It is possible to digest the acceptor  
15 vector with two different enzymes for directional cloning, however the forward and  
reverse PCR primers used to create the amplified transfer DNA must be designed to  
30 generate the necessary complementary overhangs.

The supercoiled DNA of the acceptor vector is digested with 120 units of  
35 *EcoR I* (New England BioLabs, Beverly, MA) for 3 hours under conditions specified  
by the supplier, extracted with an equal volume of phenol/chloroform/isoamyl alcohol  
20 (25:24:1), ethanol precipitated, and washed with 500 $\mu$ l of 80% ethanol (Ausubel et.  
al, Section 2.1). Ends of the DNA are dephosphorylated by treating with calf  
40 intestinal alkaline phosphatase (CIP; New England BioLabs, Beverly, MA) according  
to protocol specified by the supplier, then the DNA is extracted with  
25 phenol/chloroform/isoamyl alcohol (25:24:1), ethanol precipitated, washed with 80%  
45 ethanol, and resuspended in 1000 $\mu$ l of TE buffer, pH 8.

**C. DNA Sequence Transfer:**

**Cloning the PCR amplified product into the acceptor vector:** A combination of 4 $\mu$ l (40ng) of the topoisomerase treated PCR product (400bp - 2000bp) and 1 $\mu$ l (30ng) of the prepared acceptor vector is prepared and incubated at 25°C for 5 min., the reaction is placed on ice, and then 1 $\mu$ l of the combination is transformed into competent *E. coli* using either chemical transformation or electroporation techniques (Ausubel et. al, Section 1.8). Cells containing the acceptor vectors plus transfer sequence are selected by plating on antibiotic media requiring a resistance marker specific to the acceptor vector. Plates are incubated at 37°C for 12-18 hrs. and resulting colonies are screened by miniprep and restriction digest (Ausubel et. al, Sections 1.6 and 3.1) to identify acceptor vector clones containing the desired transfer sequence.

While the foregoing has been with reference to particular embodiments of the invention, it will be appreciated by those skilled in the art that changes in these embodiments may be made without departing from the principles and spirit of the invention, the scope of which is defined by the appended claims.

## Claims

5

10

15

20

25

30

35

40

45

50

55

That which is claimed is:

1. A cell-free subcloning system comprising:
  - a donor vector comprising a transfer sequence flanked by site-specific recombination sequences,
  - an acceptor vector comprising a site-specific recombination sequence that matches the site-specific recombination sequences of the donor vector, and
  - a site-specific recombinase capable of recognizing the site-specific recombination sequence.
2. A cell-free subcloning system according to claim 1 wherein the site-specific recombination sequence is recognized by a type I topoisomerase.
3. A cell-free subcloning system according to claim 1 wherein the site-specific recombination sequence is recognized by vaccinia DNA topoisomerase, Cre, Flp, HK022 integrase or lambda integrase.
4. A cell-free subcloning system according to claim 1 wherein the site-specific recombination sequences is identical in the donor and acceptor vectors.
5. A cell-free subcloning system according to claim 1 wherein the site specific recombination sequence is loxP, loxP511, loxB, loxC2, loxL, loxR, loxΔ117, FRT, Dif, and Att.
6. A cell-free subcloning system according to claim 3 wherein the site-specific recombination sequence is 5'-(C/T)CCTT↓, (SEQ ID NO: 1), 5'-ATAACTTCGTATA GCATACAT TATACGAAGTTAT-, (SEQ ID NO: 4), 5'-GAAGTTCCTATAC TTCTAGAA GAATAGGAACTTC, (SEQ ID NO: 7), 5'-CAAGTT, (SEQ ID NO: 12), or 5'-AACCTT, SEQ ID NO: 13).

5

7. A cell-free subcloning system according to claim 1 wherein the transfer sequence is an EST fragment, a gene sequence, or a coding sequence.

10

8. A cell-free subcloning system according to claim 1 wherein the donor vector and/or the acceptor vector additionally comprise one or more nucleic acid sequences selected from a promoter-enhancer sequence, a selection marker sequence, an origin of replication, or a fusion protein producing sequence.

15

9. A cell-free subcloning system according to claim 6 wherein the fusion protein producing sequence comprises an epitope-tag encoding sequence, an affinity purification-tag encoding sequence, or a functional protein encoding sequence.

20

10. A method of rapidly subcloning a nucleic acid sequence, said method comprising contacting a site-specific recombinase and a cell-free solution comprising a donor vector comprising a transfer sequence flanked by a site-specific recombination sequence recognized by the recombinase, and an acceptor vector comprising at least one site-specific recombination sequence recognized by the recombinase, under conditions suitable to promote the transfer of the transfer sequence from the donor vector to the acceptor vector.

25

30

11. A method according to claim 8 wherein each site-specific recombination sequence is recognized by a type I topoisomerase.

35

12. A method according to claim 8 wherein the site specific recombination sequences are identical.

20

40

13. A method according to claim 8 wherein the site-specific recombination sequence is recognized by vaccinia DNA topoisomerase, Cre, Flp, HK022 integrase or lambda integrase.

45

50

55

5 14. A method according to claim 8 wherein the site-specific recombination  
sequence is 5'-(C/T)CCTT↓, (SEQ ID NO: 1),  
10 5'-ATAACTTCGTATA GCATACAT TATACGAAGTTAT-, (SEQ ID NO: 4),  
5'-GAAGTTCCTATAC TTCTAGAA GAATAGGAACTTC, (SEQ ID NO: 7),  
5 5'-CAAGTT, (SEQ ID NO: 12), or  
5'-AACCTT, (SEQ ID NO: 13).

15 15. A method according to claim 8 wherein the transfer sequence is an  
EST fragment, a gene sequence, or a coding sequence.

20 16. A method according to claim 8 wherein the donor vector and/or the  
10 acceptor vector additionally comprise one or more nucleic acid sequences selected  
from a promoter-enhancer sequence, a selection marker sequence, an origin of  
replication, or a fusion protein producing sequence.

25 17. A method according to claim 13 wherein the fusion protein producing  
sequence comprises an epitope-tag encoding sequence, an affinity purification-tag  
15 encoding sequence, or a functional protein encoding sequence.

30 18. A subcloning kit comprising

one or more vectors, each vector comprising a site-specific  
35 recombination sequence and one or more additional nucleic acid sequences,  
wherein each vector in the kit comprises the same site-specific recombination  
20 sequence, and

40 a site-specific recombinase that recognizes the site-specific  
recombination sequence in each vector.

45 19. A subcloning kit according to claim 15 wherein the site-specific  
recombination sequence is recognized by a type I topoisomerase.

25 20. A subcloning kit according to claim 15 wherein the site-specific  
recombination sequences are identical in the vectors.

5

10

15

20

25

30

35

40

45

50

55



5

21. A subcloning kit according to claim 15 wherein the site-specific recombination sequence is recognized by a type I topoisomerase.

10

22. A subcloning kit according to claim 15 wherein the site-specific recombination sequence is recognized by vaccinia DNA topoisomerase, Cre, Flp, HK022 integrase or lambda integrase.

15

23. A subcloning kit according to claim 15 wherein the site-specific recombination sequence is 5'-(C/T)CCTT↓, SEQ ID NO: 1), (5'-ATAACTTCGTATA GCATACAT TATACGAAGTTAT-, SEQ ID NO: 4), (5'-GAAGTTCCTATAC TTCTAGAA GAATAGGAACTTC, SEQ ID NO: 7), 5'-CAAGTT, SEQ ID NO: 12), or (5'-AACCTT, SEQ ID NO: 13).

20

25

24. A subcloning kit according to claim 15 wherein the additional nucleic acid sequences are selected from a promoter-enhancer sequence, a selection marker sequence, an origin of replication, or a fusion protein producing sequence.

30

25. A subcloning kit according to claim 19 wherein the fusion protein producing sequence comprises an epitope-tag encoding sequence, an affinity purification-tag encoding sequence, a functional protein encoding sequence, or a proteolytic cleavage recognition sequence.

35

26. A kit comprising

40

at least one donor vector comprising at least one site specific recombination sequence, a transfer sequence, and a first selectable marker, and

45

at least one acceptor vector comprising at least one site specific recombination sequence, a lethal gene and a second selectable marker.

50

55

1/4

Donor

DoubleTOPO  
 GAGGATTGGG CCTT [GAT] AAGGAAGAC AATAGCAGC ATGCTGGGA TCGGTGGG TCTATGGCTT  
 CTCCTAACC GGA CTAG TTCCCTTCTG TTATCGTCCG TACGACCCCT ACGCCACCG AGATACCGAA  
 CTGAGGCGGA AGAACAGC TGGGGCTCTA GGGGGTATCC CCACGCGCCC TGTAGCGGCG CATTAAGCGC  
 GACTCCGCCT TTCTTGGTCG ACCCGAGAT CCCCATAGG GGTGCGCGG ACATCGCCG GTAATTCCG  
 GCGGGGTGTG GTGTTACGC GCAGCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCG TCCTTTCCGT  
 CCGCCACAC CACCAATGCG CGTCGCACTG GCGATGTAA CCGTCGCGG ATCGCGGCG AGGAAAGCGA  
 TTCCTCCCTT CCTTCTCGC CACGTTGCG CACGTTCCCG GTCAAGCTCT AAATCGGGG ATCCCTTTAG  
 AAGAAGGGA GGAAGAGCG GTGCAAGCG CCGAAGGG CAGTTCGAGA TTAGCCCC TAGGGAATC  
 GGTCCGATT TAGTGCTTA CGGCACCTCG ACCCAAAA ACTTGATTAG GGTGATGTT CACGTAGTGG  
 CCAAGGCTAA ATCAGGAAT GCGTGGAGC TGGGGTTTT TGAACATAAT CCACTACCA GTGCATCACC  
 GCCATCGCC TGATAGACGG TTTTTCGCCC TTGACGTTG GAGTCCACGT TCTTTAATAG TGGACTCTTG  
 CGGTAGCGG ACTATCTGCC AAAAAGCGG AACTGCAAC CTCAGGTGCA AGAATATC ACCTGAGAAC  
 DoubleTOPO  
 TTCCAAACTG GAACAACACT CAACCCCTT [CTAG] AAGGCTATC TCGGTCT  
 AAGGTTGAC CTTGTTGTA GTTGGGGAA GATC TTCCCGATAG AGCCAGA

Figure 1A

2/4

Acceptor

	Double TOPO	Double TOPO
AAAAAGCTCC	CCCTTGATCAAGGG	CGGGAGCTTG
<u>TTTTTCGAGG</u>	<u>GGGAAC TAGTCC</u>	<u>GCCCTCGAAC</u>
		GGGAAGATCTTCCC
		<u>TATATCCATT</u>
		<u>ATATAGGTAA</u>

Figure 1B

## Recombint

DoubletOP0

AAAAAGCTCC CCCTTGATC)AAGGAAGAC AATAGCAGGC ATGCTGGGA TCGGTGGG TCTATGGCTT  
TTTTTCGAGG GGAAGTAG TTCCCTTCTG TTATCGTCCG TAGGACCCCT ACGCCACCG AGATACCGAA  
 CTGAGCGGA AGAACCAGC TGGGGCTCTA GGGGGTATCC CCACGGCCG TGTAGCGCG CATTAAGCGC  
 GACTCCGCTT TTCTTGGTCG ACCCCGAGAT CCCCCATAGG GGTGCGCGG ACATCGCCG GTAATTGCGG  
 GCGGGGTGTG GTGGTTACGC GCAGCGTGAC CGCTACACTT GCCAGCGCC TAGCGCCCG TCCTTTGCTT  
 CCGCCACAC CACCAATGCG CGTCGCACTG GCGATGTGAA CCGTCGCGG ATCGCGCGG AGAAAGCGA  
 TTCTTCCCTT CCTTCTCGC CACGTTCCG GGTTCCTCC GTCAAGCTCT AAATCGGGG ATCCCTTTAG  
 AAGAAGGGAA GGAAGAGCG GTGCAAGCG CGAAAGGG CAGTTCGAGA TTAGCCCC TAGGAAATC  
 GGTTCGGATT TAGTGCTTTA CGGCACCTCG ACCCCAAAA ACTTGATTAG GGTGATGGT CACGTAGTGG  
 CCAAGGCTAA ATCAGGAAT GCCGTGGAGC TGGGGTTTT TGAATAATC CCACTACCA GTGCATCACC  
 GCCATCGCC TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCAGT TCTTTAATAG TGGACTCTTG  
 CCGTAGCGG ACTATCTGCC AAAAGCGG AACTGCAAC CTCAGGTGCA AGAAATTATC ACCTGAGAAC

DoubletOP0

TTCCAAACTG GAACAACACT CAACCCCTT(CTAG)AGGG TATATCCATT  
 AAGGTTTGAC CTTGTTGTGA GTTGGGGGA GATCTTCCC ATATAGGTAA

Figure 1C

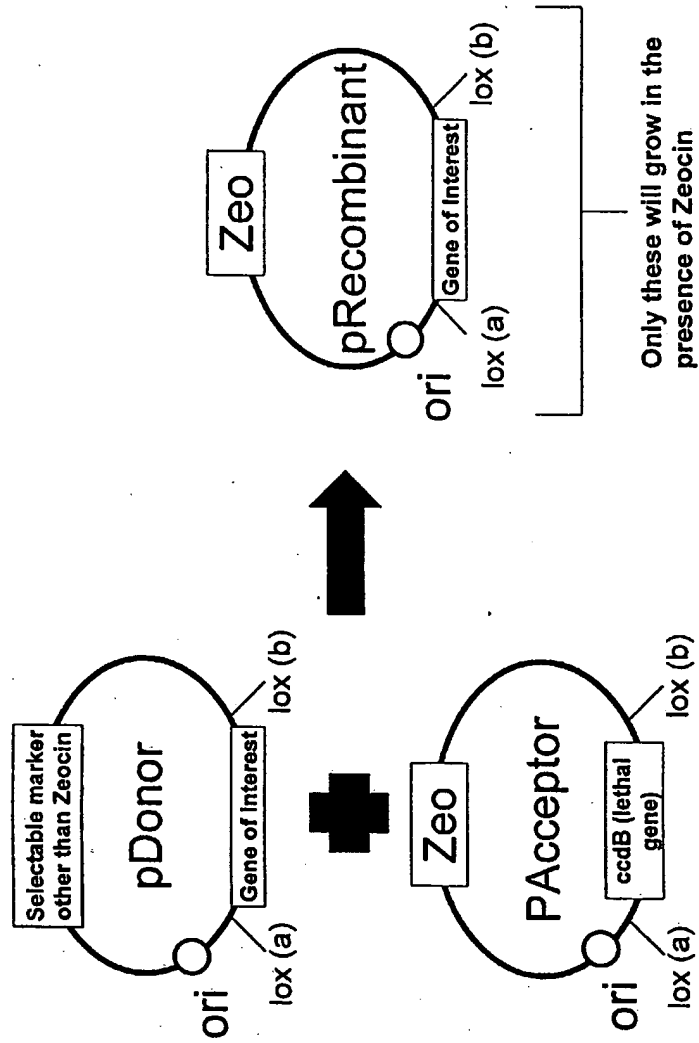


Figure 2

## SEQUENCE LISTING

<110> Miles, David J.  
Turner, Lyle C.  
Marcil, Bob  
McConnell, Gina

<120> System for the Rapid Manipulation of  
Nucleic Acid Sequences

<130> INVIT1160

<160> 16

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 5

<212> DNA

<213> Artificial Sequence

<220>

<400> 1

ycctt

5

<210> 2

<211> 10

<212> DNA

<213> Artificial Sequence

<220>

<400> 2

nnnnaagggc

10

<210> 3

<211> 10

<212> DNA

<213> Artificial Sequence

<220>

<400> 3

gcccttnnnn

10

<210> 4

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<400> 4

ataacttctg atagcataca ttatacgaag ttat

34

<210> 5  
<211> 34  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
  
<400> 5  
ataacttcgt atagtataca ttatacgaag ttat 34  
  
<210> 6  
<211> 34  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
  
<400> 6  
acaacttcgt ataatgtatg ctatacgaag ttat 34  
  
<210> 7  
<211> 34  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
  
<400> 7  
gaagttccta ttctctagaa agtataggaa cttc 34  
  
<210> 8  
<211> 14  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
  
<400> 8  
ttcaatyaag gtaa 14  
  
<210> 9  
<211> 10  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
  
<400> 9  
gcccgargg 10  
  
<210> 10  
<211> 12  
<212> DNA  
<213> Artificial Sequence

<220>  
<400> 10  
gaagcagtgg ta 12  
<210> 11  
<211> 3  
<212> DNA  
<213> Artificial Sequence  
<220>  
<400> 11  
ttg 3  
<210> 12  
<211> 6  
<212> DNA  
<213> Artificial Sequence  
<220>  
<400> 12  
caagtt 6  
<210> 13  
<211> 6  
<212> DNA  
<213> Artificial Sequence  
<220>  
<400> 13  
aacctt 6  
<210> 14  
<211> 5  
<212> DNA  
<213> Artificial Sequence  
<220>  
<400> 14  
aaggg 5  
<210> 15  
<211> 7  
<212> DNA  
<213> Artificial Sequence  
<220>  
<400> 15  
cgaaggg 7



WO 00/12687

PCT/US99/19413

4

<210> 16  
<211> 12  
<212> DNA  
<213> Artificial Sequence

<220>

<400> 16  
cccttgcaag gg

12

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/19413

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12N 15/00, 15/64, 15/66; C12P 19/34, 21/06; C12Q 1/68

US CL : 435/6, 69.1, 91.41, 91.42, 91.5, 91.52, 320.1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 69.1, 91.41, 91.42, 91.5, 91.52, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NONEElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	US 5,888,732 A (HARTLEY et al.) 30 March 1999, see entire document.	1-26
Y	US 5,766,891 A (S. SHUMAN) 16 June 1998, see entire document.	1-26
Y	HOESS et al. P1 site-specific recombination: Nucleotide sequence of the recombining sites. Proceedings of the National Academy of Sciences, USA. June 1982, Vol. 79, pages 3398-3402, see entire document.	1-26
Y	ABREMSKI et al. Bacteriophage P1 Cre-loxP site-specific recombination. The Journal of Biological Chemistry. 05 January 1986, Vol. 261, No. 1, pages 391-396, see entire document.	1-26

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

10 DECEMBER 1999

Date of mailing of the international search report

27 JAN 2000

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

IREM YUCEL

Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/19413

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NASH et al. Heteroduplex substrates for bacteriophage lambda site-specific recombination: cleavage and strand transfer products. The EMBO Journal. 1989, Vol. 8, No. 11, pages 3523-3533, see entire document.	1-26
Y	SHUMAN, S. Recombination mediated by vaccinia virus DNA topoisomerase I in <i>Escherichia coli</i> is sequence specific. Proceedings of the National Academy of Sciences, U.S.A. November 1991, Vol. 88, pages 10104-10108, see entire document.	1-26

**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/US99/19413

**B. FIELDS SEARCHED**

Electronic data bases consulted (Name of data base and where practicable terms used):

USPATENTS, DERWENT, STN-CAPLUS, DIALOG, MEDLINE, BIOSIS, SCISEARCH

terms: cell-free, in vitro, recombinase?, clon?, subclon?, Cre, lox?, att?, site, specific, integrase?, topoisomerase?, lethal, toxic, fusion, protein?, polypeptide?, kif?, donor?